Supporting Information:

Distribution of extracellular flavins in a coastal marine basin and their relationship to redox gradients and microbial community members

Danielle R. Monteverde[a]†*, Jason B. Sylvan[b], Christopher Suffridge[c]†*, J. Jotautas Baronas[a]†*, Erin Fichot[c], Jed Fuhrman[c], William Berelson[a], Sergio A. Sañudo-Wilhelmy[a,c]

[a] Department of Earth Sciences, University of Southern California, Los Angeles, California, USA

[b] Department of Oceanography, Texas A&M University, College Station, Texas, USA

[c] Department of Biological Sciences, University of Southern California, Los Angeles, California, USA

#Address correspondence to Danielle Monteverde, dmonteverde@caltech.edu

P: 626-395-6123, F: 626-395-1995

†Present addresses:

DRM, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California, USA

CS, Department of Microbiology, Oregon State University, Corvallis, Oregon, USA

JJB, Department of Earth Sciences, University of Cambridge, Cambridge, United Kingdom

This supporting information contains a 20-page document which includes 9 figures, 8 tables, a brief description of methods, and this cover page.

1. Background Information

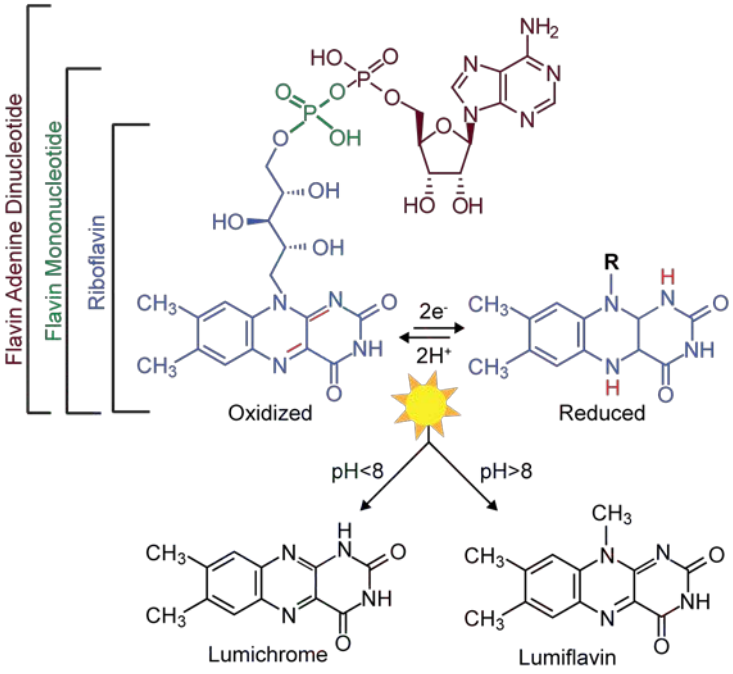

Figure S1. Flavin structures. Riboflavin (RF, in blue) is the precursor to flavin mononucleotide (FMN, in green) and flavin adenine dinucleotide (FAD, in red). All three molecules are capable of single or double electron and hydrogen transfer (semiquinone not shown). Under light conditions the RF, FMN, and FAD will photodegrade and lose the ribityl side chain to form lumichrome or lumiflavin depending on pH.

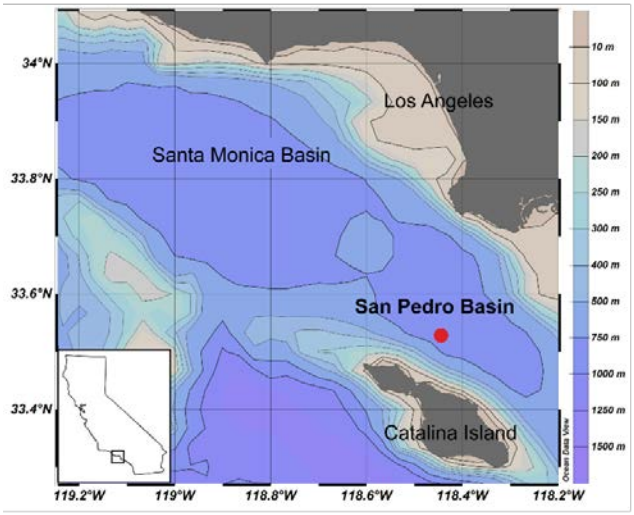

Figure S2. San Pedro Ocean Time Series (SPOT), 33° 33’ N, 118° 24’W.

2. Analytical Protocol

2.1. Instrument Settings

A Thermo high-performance liquid chromatography/tandem triple quadrupole mass spectrometer (LC/MS) instrument (Waltham, MA, USA) was optimized via a direct infusion of each standard in methanol to find the relevant positive/negative ion mode, as well as parent and product ions of the flavin analytes (see Table S1). The mass spectrometer was operated in the positive ion mode and analytes were detected in the selected reaction monitoring (SRM) mode with the instrument parameters listed in Table S2.

Table S1. LC/MS Flavin Optimization Conditions

Standard	M_r (purity)	RT (minutes)	Parent Ion $[M+H]^+$	Product Ions	CE (V)
Riboflavin (RF)	376.38 (≥98%)	7.35	377.15	172.01 197.97 242.94	34 37 22
Riboflavin – dioxopyrimidine-$^{13}C_4$, $^{15}N_2$ (Heavy RF)	382.32 (97%)	7.35	383.1	174.9 202.3 248.9	35 36 22
Riboflavin 5'-monophosphate sodium salt hydrate (FMN-Na)	478.33 (≥95%)	8.25	457.4	242.9 359.3 438.7	30 23 13
Lumichrome (LC)	242.23	8.40	242.98	103.09 172.03 198.00	36 22 23
Lumiflavine (LF)	256.26	7.70	**257.1**	159.01 185.99 213.07	10 22 17

Abbreviations: M_r, relative molecular weight; RT, retention time; CE, collision energy.

Table S2. LC/MS Instrument Parameters

Spray voltage (V)	4000
Sheath gas (N_2) pressure	30
Ion sweep gas pressure	3.0
Auxiliary gas pressure	5
Capillary temperature (°C)	269
Capillary offset (V)	35
Scan width (m/z)	0.1
Tube lens offset (V)	87
Scan time (s)	0.1
Collision gas pressure (mTorr)	2.1

2.2.Flavin Calibration

Standards were prepared fresh daily from a stock standard that was prepared every 2 weeks and stored protected from light at -20ºC except RF which was stored at 4ºC as instructed by the manufacturer. Standards and external calibration curves were run daily (see Figure S3). Calibration curves included the internal standard, which was used to adjust the daily Response Factor value to account for instrumental drift.

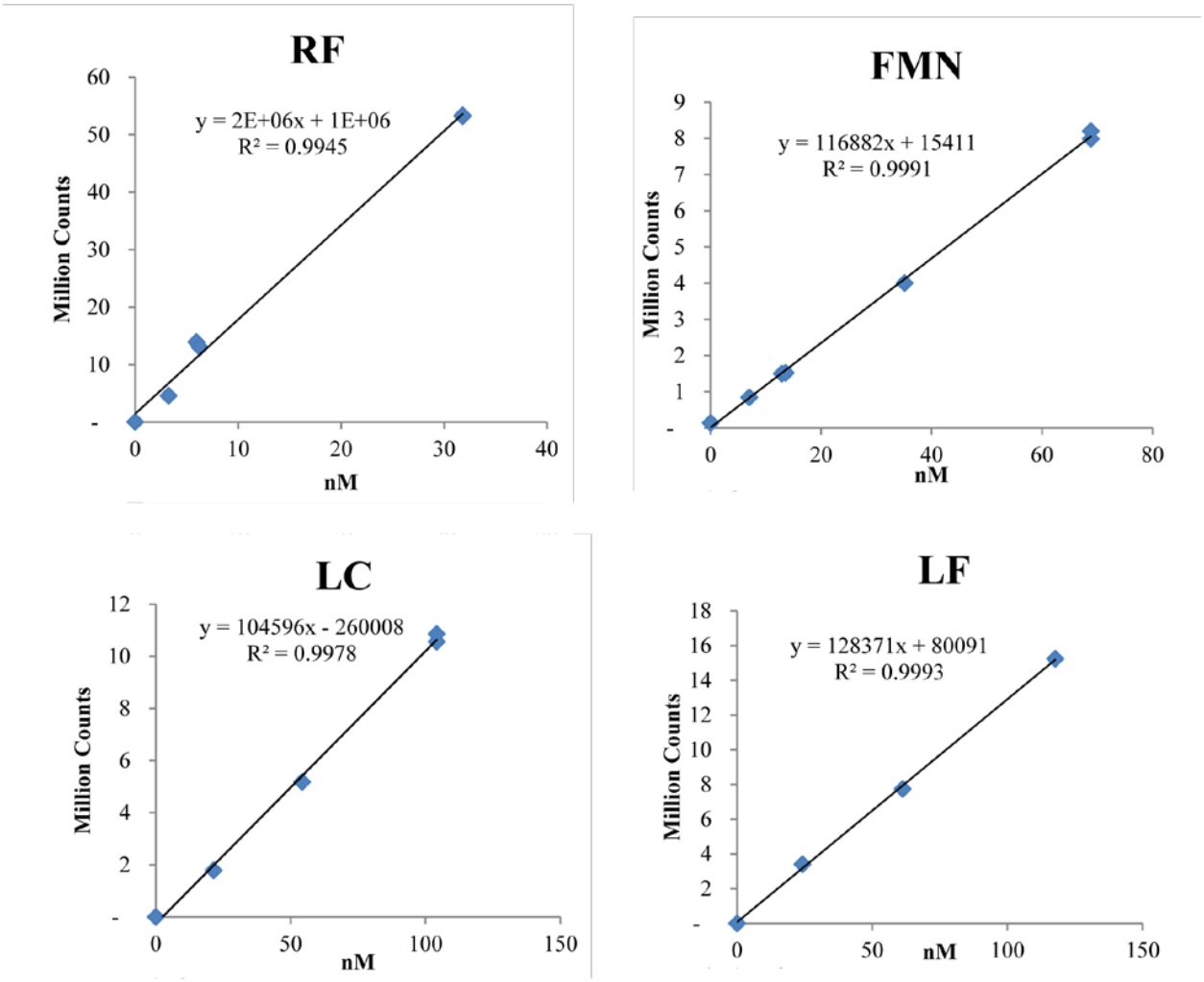

Figure S3. Example LC/MS Calibration Curves for flavin analytes riboflavin, flavin mononucleotide, lumichrome, and lumiflavin.

2.3.Response Factor

An instrumental response factor was calculated to establish a relationship between the internal standard and each individual analyte. This response factor was calculated daily to account for any variations in instrument sensitivity as previously described. The response factor was calculated with the following equation:

$$\text{Response Factor} = \frac{A_{IS}}{C_{IS}} \frac{C_{vit}}{A_{vit}}$$

Where A_{IS} is the signal area or counts of the internal standard, C_{IS} is the known concentration of the internal standard, C_{vit} is the concentration of the measured analyte and A_{vit} is the signal area of the measured analyte. Once the response factor was determined sample quantification was achieved by rearranging the equation to solve for C_{vit}. Response factors are listed in Table S3.

2.4.Limits of Detection

Procedural blanks and spikes were run for each sample set. RF and FMN were included in the procedural spike but LC and LF were omitted. All analytes were included in method development spikes and demonstrated high recoveries (>80%).

The analytical LOD was calculated as either 3 times the standard deviation of the procedural blank or 3.3 times the standard error of the y-intercept of the calibration curve. The LOD reported was the lowest of these two numbers. Note that this analytical LOD is for the LC/MS and not the full preconcentration procedure. The procedural LOD will vary based on the volume of water concentrated but will generally be at least two orders of magnitude lower. For instance, water column samples consisted of 1-L samples concentrated down to ~400 µL resulting in a concentration factor of 2500.

Table S3. Analyte detection conditions

Analyte	% Recovery	Analytical LOD (nM)	Procedural LOD (pM)	Slope*	R^2	Response Factor*
Riboflavin	74-101	0.56	0.22	1.6 E6	0.9945	0.029
Flavin Mononucleotide	77-97	1.60	0.64	1.2 E5	0.9991	0.387
Lumichrome		2.83	1.13	1.0 E5	0.9978	0.382
Lumiflavin	86-110	1.19	0.47	1.3 E5	0.9993	0.415

*Slope and Response Factor are measured and adjusted daily

2.5.Chloroform Extraction

A chloroform extraction was tested on flavin standards to assess any loss of signal to the desired analytes. It was found that chloroform did not affect recovery of FMN or RF (chloroform recoveries of 91-105%); however, chloroform extraction did effectively remove the degradation products LC and LF from solution (24 and 13% recovery, respectively). Figure S2 shows the results of the chloroform extraction test. The liquid phase extraction (LPE) procedure involved addition of 50 µL of chloroform to the preconcentrated sample (1:5 v chloroform/v sample) followed by 2 minutes of agitation on a vortexer, centrifugation for 3 minutes at 5000 rpm, and removal of the denser chloroform fraction. A chloroform extraction was not performed on the water column samples as matrix effects did not inhibit quantification and therefore degradation products LC and LF were also reported.

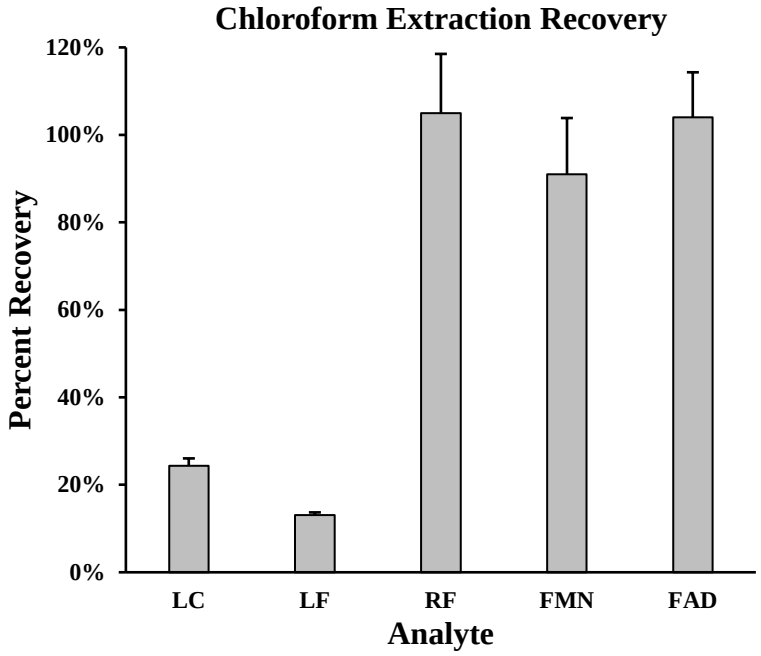

Figure S4. Recovery of flavin standards following a chloroform extraction. Error bars represent a triple injection of the sample.

3. Results

3.1.Flavin and Geochemical Concentrations

Table S4. Geochemical Concentrations

Core A Depth (cm)	Dissolved Fe (µM)	Stdev	Dissolved Mn (nM)	Stdev	Core C Depth (cm)	SO_4^{2-} (mM)
0.5	219.61	13.31	46.58	3.10	0	26.5
3.5	207.12	19.47	98.74	6.02	1	26.5
5.5	131.59	3.04	119.71	1.55	3	24.4
10.5	134.25	16.45	184.95	0.67	5	26.1
15.5	98.31	13.25	209.74	0.46	9	26.3
20.5	89.74	8.04	270.66	4.29	16	26.0
25.5	64.33	4.04	333.62	7.16	24	24.5
30.5	50.75	0.18	348.44	3.12	29	24.9
36.0	41.70	3.62	361.51	7.55	33	23.5
42.5	18.01	1.57	475.77	15.28		

Table S5. Dissolved Flavin Water Column Concentrations

Depth	RF		FMN		Lumichrome		Lumiflavin	
(m)	(pM)	Stdev	(pM)	Stdev	(pM)	Stdev	(pM)	Stdev
10	23.91	4.34	122.87	32.90	137.23	22.99	24.99	6.93
50	21.77	10.38	121.41	44.96	9.75	5.01	4.77	2.06
100	20.06	0.72	104.69	1.81	9.04	2.19	6.51	2.03
300	23.03	0.43	129.45	0.23	9.43	0.95	4.68	0.55
500	16.43	4.37	93.21	2.10	7.58	1.35	3.22	1.21
700	14.11	0.57	95.99	7.41	4.68	0.17	2.77	0.72
800	18.81	1.11	94.04	7.09	6.33	0.22	2.37	0.14
870	40.27	7.97	295.88	20.28	10.53	3.39	3.60	0.34

Table S6. Dissolved Flavin Pore Water Concentrations

Core A					Core B				
Depth	RF		FMN		Depth	RF		FMN	
(cm)	(pM)	Stdev	(pM)	Stdev	(cm)	(pM)	Stdev	(pM)	Stdev
0.0	62.94	20.20	613.02	176.53	0.0	10.80	0.14	253.90	19.89
0.5	88.55	10.94	737.63	8.62	0.5	25.89	2.22	399.81	17.06
3.5	93.42	4.92	1181.65	59.61	3.5	21.56	2.63	573.86	41.16
5.5	82.83	5.65	894.17	14.63	6.5	55.47	11.42	604.86	48.60
10.5	114.70	16.68	1087.06	56.11	9.5	43.15	4.28	391.55	41.83
15.5	99.15	12.97	1319.80	241.02	15.5	57.10	15.11	596.50	35.29
20.5	128.73	31.38	1424.94	102.66	18.5	73.87	9.95	749.82	47.24
25.5	115.94	22.77	1286.93	262.00	22.5	65.26	9.34	786.43	75.64
30.5	165.81	5.79	1880.31	115.08	28.5	86.40	3.50	1116.58	71.69
36.0	210.72	20.23	1728.52	108.81	34.0	NA	NA	1019.47	21.95
42.5	193.14	42.51	2072.01	223.34	40.0	89.57	8.75	1086.93	34.68

3.2.Flavin Statistical Comparisons

Basic statistical correlation analyses were computed in the software program R. Core A was used as the primary dataset since this was the core that had measurements of both vitamins and dissolved Fe. Statistically significant (>95% confidence) correlations were found between all combinations of RF, FMN, dissolved Fe, and Depth (Figure S5). A Pearson correlation was also calculated for all of the FMN and RF samples (water column and porewaters from both cores) and found a strong, statistically significant correlation (see Figure S6).

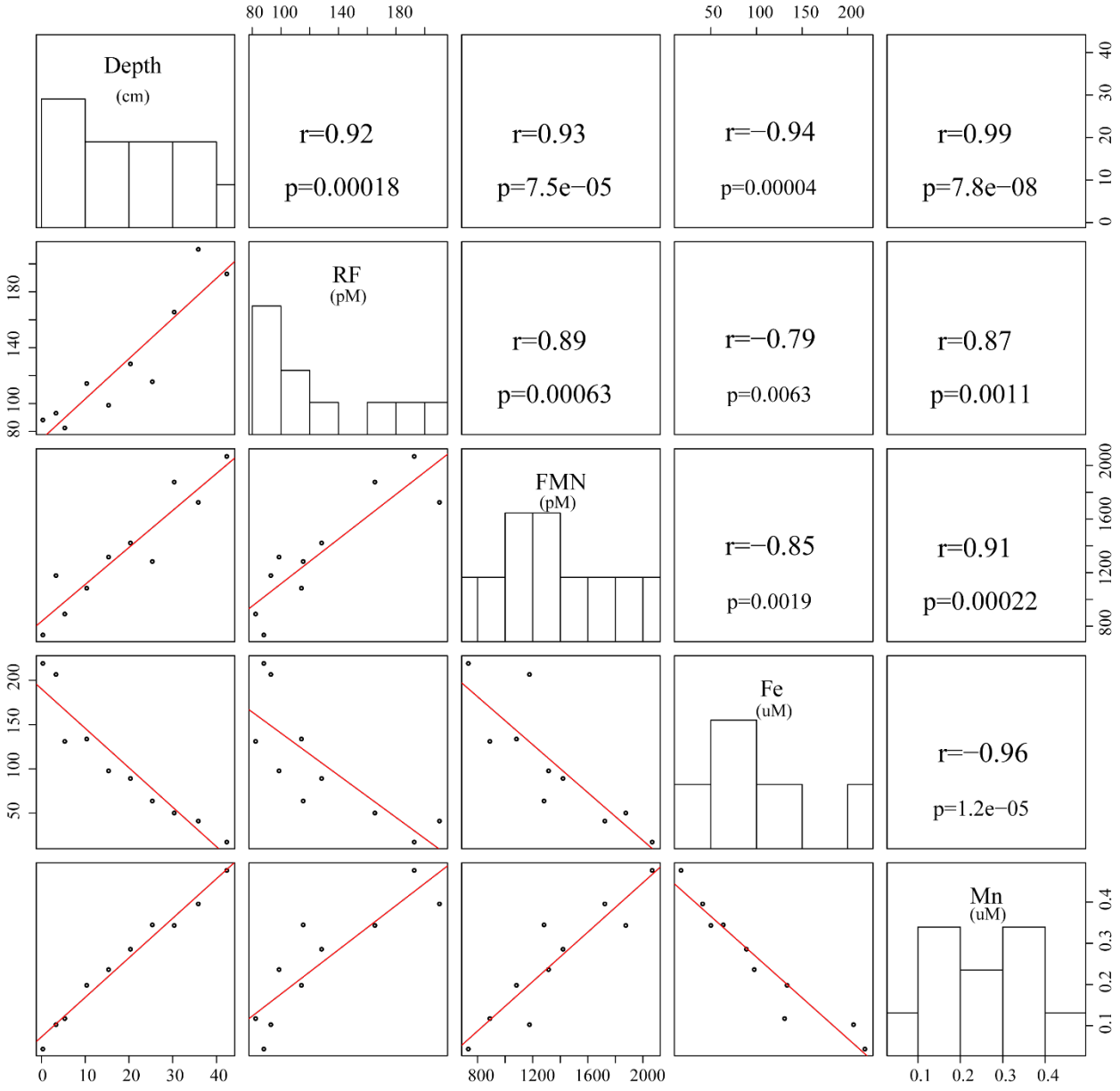

Figure S5. Statistical analyses of pore water samples for Core A**.** The figure shows the histogram distribution of the samples for Core A which divide the plotted correlations on the bottom left from the Pearson correlation coefficient and probability for each correlation in the upper right. Any value of $p<0.05$ is statistically significant at the 95% confidence interval.

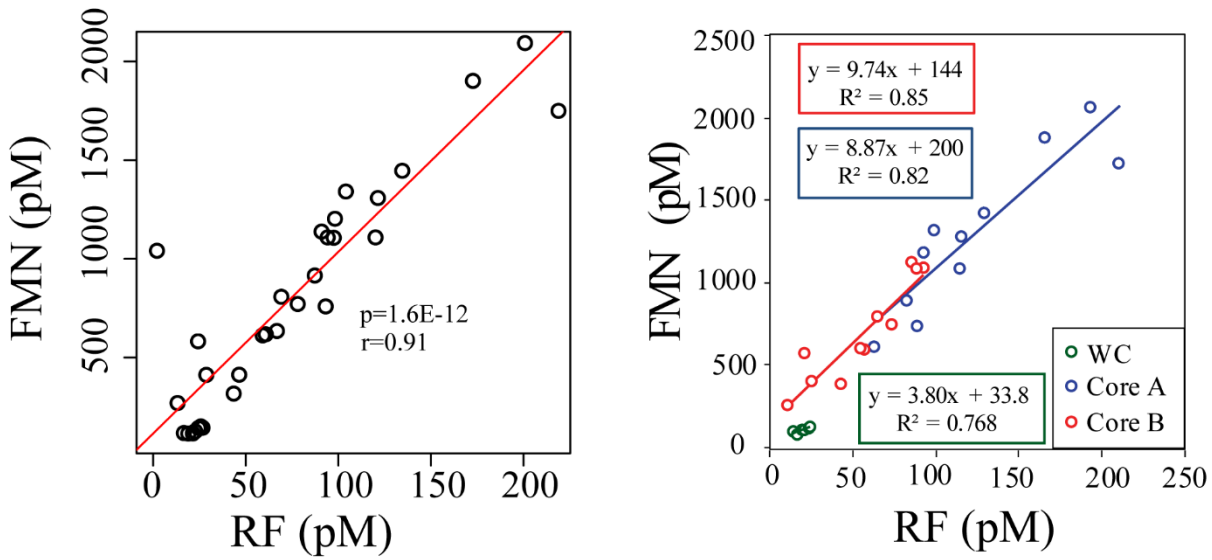

Figure S6. Pearson Correlation between FMN and RF in water column and pore water samples for all water column and pore water samples on the Left and for individual sample sets (WC = water column, Core A and B = pore water) on the Right.

4. Microbial Community Analysis

Cell enumeration

Direct cell counts were performed to quantify microbial abundance using an adaptation of a previously published method to separate cells from sediment grains.[2] Briefly, 1 mL of cell suspension was added to a mix of 2.2 ml 2.5% NaCl, 400 μL detergent mix (100 mM EDTA, 100 mM sodium pyrophosphate, 1% [v/v] tween-80) and 400 μL methanol and shaken at 500 rpm for 10 minutes on a Troemner Digital Multi-tube Vortex Mixer (Thorofare, NJ). The sediment was then pelleted by centrifugation at 3000 x G for 5 minutes and 0.4-2.0 mL of supernatant filtered onto 25 mm, 0.2 μm black polycarbonate filters.[3] After staining the filter samples with DAPI (45 μM final concentration) for 5 minutes in the dark, filters were mounted on glass slides and enumerated using an epifluorescence microscope (Zeiss Axio Imager.M2) at 1000× magnification.

PCR

Duplicate extracted community DNA from the same sample used for cell enumeration was combined for polymerase chain reaction (PCR) analysis of the V4/5 region of the 16S rRNA gene. PCR was performed either in triplicate and pooled (Set 1, Table S7) or in singleton (Set 2). Each 25-μL reaction contained 0.5-ng DNA template, 0.3 mM of each primer (Eurofins, 515F: GTGYCAGCMGCCGCGGTAA; 926R CCGYCAATTYMTTTRAGTTT; 2% PVP-40, and 12 μL Master Mix (5 Prime HotMasterMix: 0.6 U taq, 54 mM KCl, 3 mM Mg^{2+}, 240 μM each dNTP). Primers were constructed with Illumina sequencing primers and adapters inline[4] with the modification of including the barcode on the forward primer after the four base pair N pad instead of including the barcode on the reverse primer. Reaction conditions included 2 min at 95°C; 30 cycles of 45 sec at 95°C, 45 sec at 50°C, and 1.5 min at 68°C; and 5 min at 68°C. Reactions were visualized by gel electrophoresis, purified (Agencourt® Ampure® XP, 1x beads), quantified (Pico-Green dsDNA Quant-iT Assay Kit), and pooled with other amplicons for sequencing.

OTU Assembly

We constructed 927,686 contigs with an average read length of 374bp distributed across 11 experimental samples, two extraction blanks and a PCR blank. After screening of sequences with ambiguous bases or long homopolymers, 553,700 unique sequences were used to generate an alignment. Further quality control measures (i.e., pre-clustering, alignment curation, and chimera removal) reduced the total library to 129,791 unique contigs. The two extraction blanks and PCR blank libraries contained 489 Operational Taxonomic Units (OTUs). Of these, the blank extraction from the first sequencing plate was found to contain all 489 OTUs, with the blank extraction from the second run having only 3 OTUs (8 sequences, total), therefore OTUs present in the blank were removed from the entire dataset if they met one of the two criteria: (1) the OTU was present in the second sequencing run of samples, or (2) the OTU was present in one or more of the first run of samples at an abundance <10X it's abundance in the blank from the same run. This resulted in a dataset with 617,623 sequences, from which singletons were removed, resulting in a final curated dataset containing 17,897 OTUs clustered from 605,012 iTags. Each library contained an average of 55,001 ± 31,902 sequences.

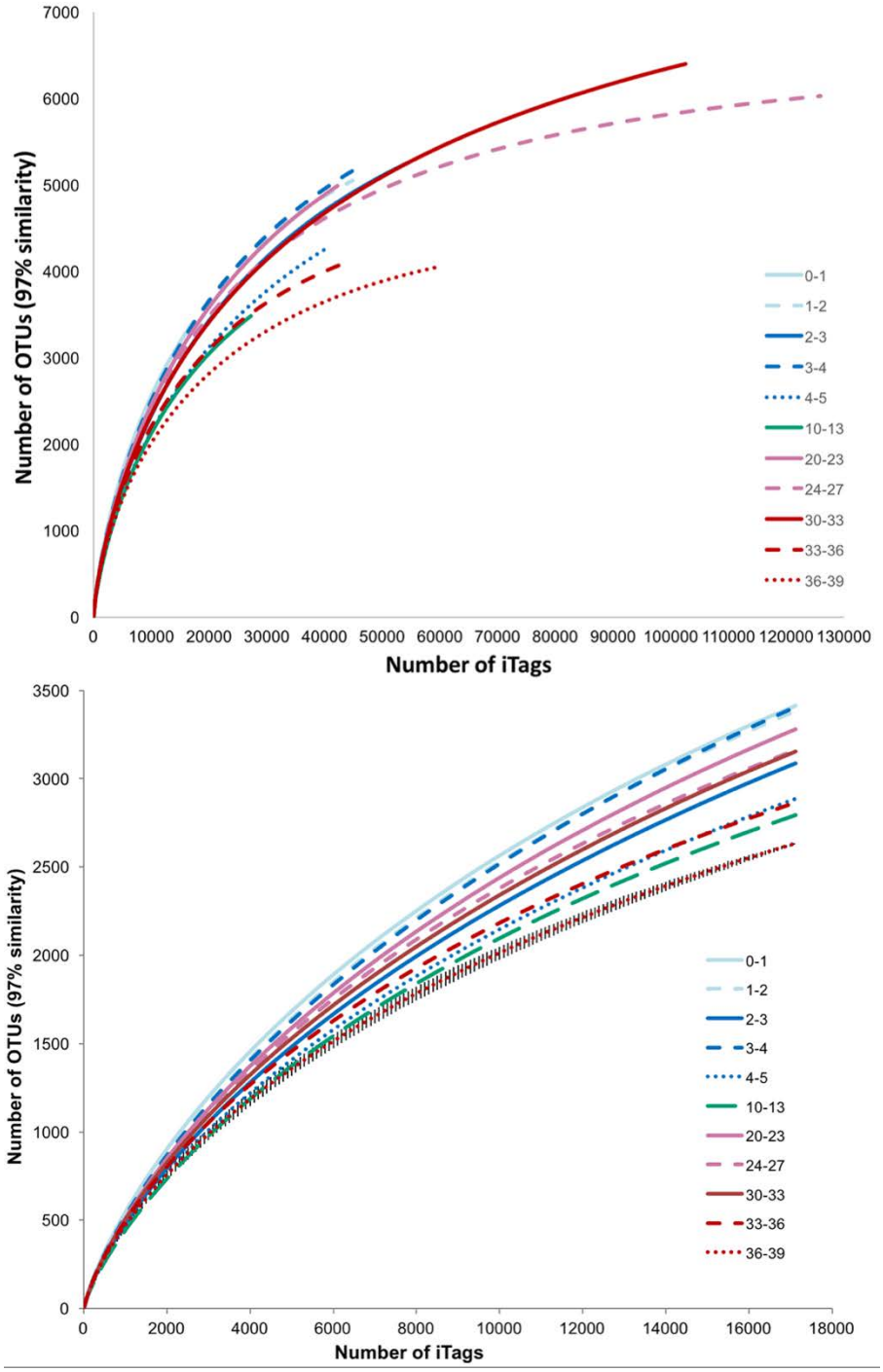

Figure S7. Rarefaction analysis for 16S rRNA amplicons. Legend numbers correspond to homogenized sample depths in centimeters. Data presented for unrarefied datasets (top) and for each sample was subsampled to 17,131 iTags, the number of reads present in the library with the fewest reads (bottom).

Table S7. Summary of data for microbiological samples. Set # reflects the sequencing run; 2X300 bp MiSeq (Set 1) and 2X250 HiSeq (Set 2). # iTags (pre-subsample) reflects final analyzed library for taxonomy, correlations and network analysis, post quality control and blank subtraction. Calculations for # OTUs, Good's coverage, and Inverse Simpson based on subsampling to 17,131 reads, the size of the smallest library.

Sample	Set #	# iTags (pre-subsample)	# OTUs	Good's Coverage	Inverse Simpson	Cell Count
Subsampled to 17,131 iTags						
0-1 cmbsf	1	17,131	3413	0.902	209.5	$5.53x10^8$
1-2 cmbsf	1	44,972	3384	0.901	188.5	$4.48x10^8$
2-3 cmbsf	1	54,483	3088	0.905	103.1	$2.69\ x10^8$
3-4 cmbsf	1	45,628	3407	0.896	128.8	$2.89\ x10^8$
4-5 cmbsf	1	40,612	2888	0.913	65.7	$1.50\ x10^8$
10-13 cmbsf	1	27, 315	2794	0.920	62.0	$1.27\ x10^8$
20-23 cmbsf	1	42,271	3281	0.901	128.1	$8.30\ x10^7$
24-27 cmbsf	2	125,946	3160	0.911	109.2	$4.20\ x10^7$
30-33 cmbsf	2	102,556	3152	0.904	135.2	$1.84\ x10^7$
33-36 cmbsf	2	44,747	2866	0.922	142.2	n.d.
36-39 cmbsf	2	59,351	2633	0.929	134.0	n.d.

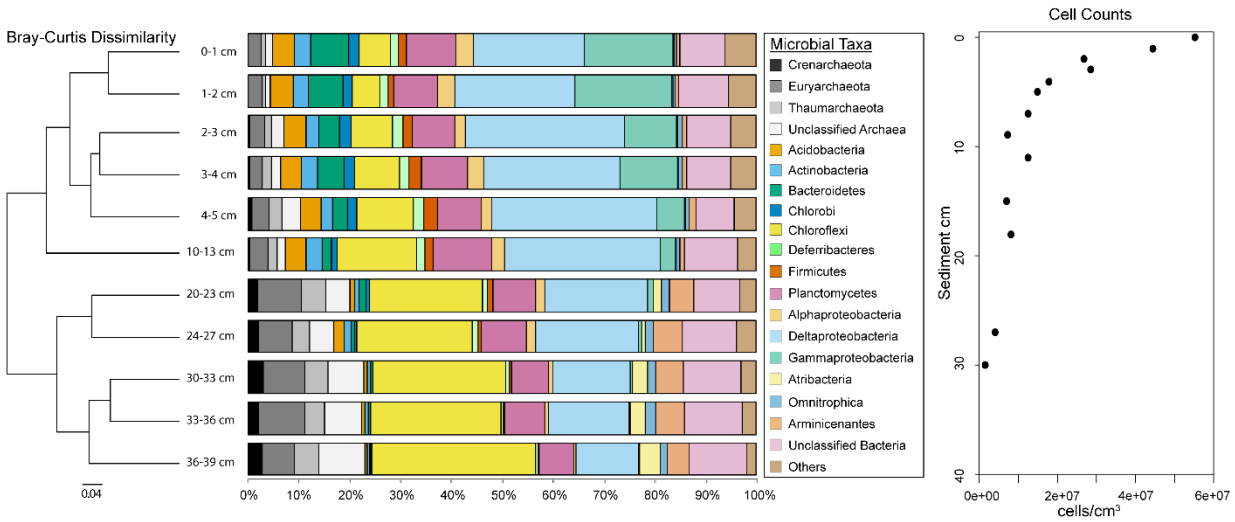

Figure S8. 16S rRNA Microbial Community Profile of SPOT sediments. Bray-Curtis dissimilarity index between sample depths shown in bars on the left. Colored bars represent relative abundance of phyla at respective sediment horizons. Phyla present in low abundance are grouped into the "Others" category. The left panel displays cell counts. Note that there is a difference in y-axis depth steps between community profile and cell count samples.

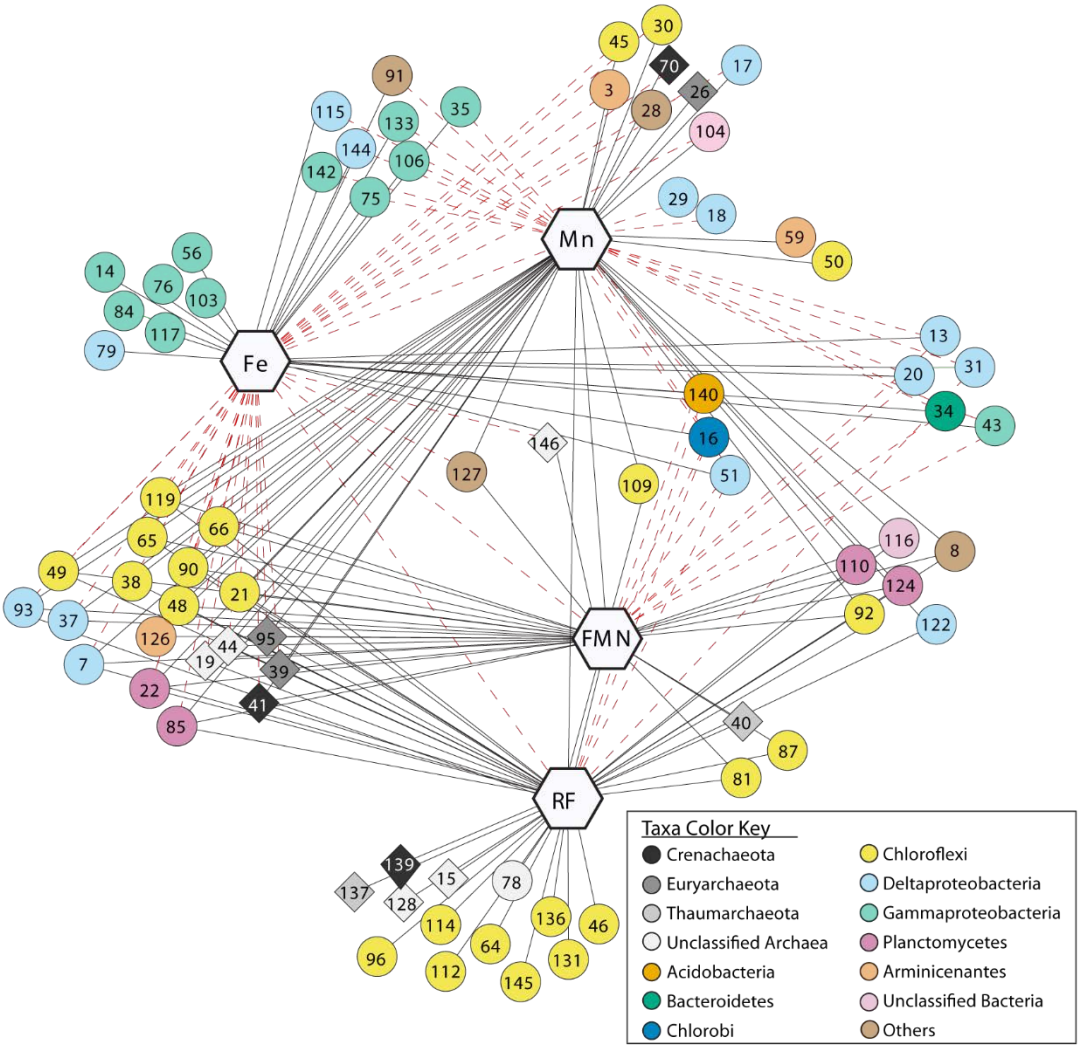

Figure S9. Network analysis of 16S rRNA OTUs with correlations to at least one of riboflavin (RF), flavin mononucleotide (FMN), dissolved Mn and/or Fe. Same data as presented in Fig. 3, but here with OTU #s labeled. Black lines indicate positive correlations, dashed red lines indicate negative correlations. Nodes are labeled with OTU #s given in Table S9 (separate excel file), with 0's before the number removed for clarity (e.g. Otu000021 in Table S9 is labeled here as "21"). Circles represent bacterial OTUs and diamonds represent archaeal OTUs.

Table S8. Correlations between lineages and flavins. Linear regression statistics for select lineages where total relative abundance of lineage was >0.10% of the dataset and r^2 >0.60

Kingdom;Phylum;Class;Order	Slope, RF	r^2, RF	Slope, FMN	r^2, FMN
Bacteria				
Chloroflexi				
Ardenticatenia	-2.650	0.710	-2.650	0.753
Dehalococcoidia				
GIF3	3.990	0.859	3.658	0.687
GIF9	0.529	0.918	0.469	0.722
vadinBA26	0.231	0.887	0.220	0.804
unclassified	0.671	0.792	0.832	0.679
S085	2.501	0.839	2.533	0.861
Proteobacteria				
Deltaproteobacteria				
Desulfarculales				
Desulfarculaceae	0.449	0.825	0.465	0.883
Desulfobacterales				
Desulfobulbaceae	-0.580	0.667	-0.603	0.722
Myxococcales				
Sandaracinaceae	-1.253	0.685	-1.297	0.734
Sh765B-TzT-29	-0.569	0.759	-0.589	0.812
Atribacteria	0.641	0.917	0.600	0.802
Omnitrophica	1.984	0.796	2.110	0.704
Arminicenantes	0.331	0.551	0.374	0.701
Archaea				
Crenarchaeota				
Thermoprotei				
unclassified	0.768	0.822	0.758	0.803
Euryarchaeota				
Thermoplasmata				
Thermoplasmatales	0.496	0.737	0.525	0.827
unclassified	1.210	0.575	1.365	0.732
Thaumarchaeota				
Misc. Crenarchaeotic Group	1.285	0.865	1.230	0.793
unclassified	4.752	0.849	4.284	0.690

References

1. Suffridge, C.; Cutter, L.; Sanudo-Wilhelmy, S., A new analytical method for direct measurement of particulate and dissolved B-vitamins and their congeners in seawater. *Front. Mar. Sci.* **2017,** *4* (11), 1-11.
2. Kallmeyer, J.; Smith, D. C.; Spivack, A. J.; D'Hondt, S., New cell extraction procedure applied to deep subsurface sediments. *Limnol. Oceanogr.: Methods* **2008,** *6* (6), 236-245.
3. Hewson, I.; Jacobson-Meyers, M. E.; Fuhrman, J. A., Diversity and biogeography of bacterial assemblages in surface sediments across the San Pedro Basin, Southern California Borderlands. *Environ. Microbiol.* **2007,** *9* (4), 923-933.
4. Bartram, A.; Lynch, M.; Stearns, J.; Moreno-Hagelsieb, G.; Neufeld, J., Generation of multimillion-sequence 16S rRNA. *Appl. Environ. Microbiol.* **2011,** *77* (11), 3846–3852.